Estimating the p-variation index of a sample function: An application to financial data set

RIMAS NORVAIŠA[1] norvaisa@ktl.mii.lt
Institute of Mathematics and Informatics, Akademijos 4, Vilnius 2600, Lithuania

DONNA MARY SALOPEK[2] dsalopek@mathstat.yorku.ca
York University, Department of Mathematics and Statistics, 4700 Keele Street, Toronto, Ontario M3J 1P3 Canada

Abstract. In this paper we apply a real analysis approach to test continuous time stochastic models of financial mathematics. Specifically, fractal dimension estimation methods are applied to statistical analysis of continuous time stochastic processes. To estimate a roughness of a sample function we modify a box-counting method typically used in estimating fractal dimension of a graph of a function. Here the roughness of a function f is defined as the infimum of numbers $p > 0$ such that f has bounded p-variation, which we call the p-variation index of f. The method is also tested on estimating the exponent $\alpha \in [1, 2]$ of a simulated symmetric α-stable process, and on estimating the Hurst exponent $H \in (0, 1)$ of a simulated fractional Brownian motion.

Keywords: estimation, p-variation index, box-counting index, financial data analysis

AMS 1991 Subject Classification: 90A20, 62M09, 60G17

1 Introduction: the p-variation index and financial modeling

In financial mathematics, a simplest continuous time model assumes that a stock price, or other financial asset, is a stochastic process $P = \{P(t)\colon 0 \le t \le T\}$ satisfying the relation

$$P(t) = 1 + (I)\int_0^t P\,dB, \qquad 0 \le t \le T, \tag{1}$$

which is the Itô integral equation with respect to a standard Brownian motion $B = \{B(t)\colon t \ge 0\}$. Equation (1) is usually written in the form of a stochastic differential equation, or simply by giving its solution $P(t) = \exp\{B(t) - t/2\}$, $0 \le t \le T$. In the financial literature this is known as the Black-Scholes-Merton stock price model. Its pertinence is backed-up by the assumption that increments of the log transform of a stock price are independent and normally distributed, which is known as the strong form of Random Walk Hypothesis. In agreement with a relaxed form or an alternative form of the Random Walk Hypothesis, the Brownian motion B in equation (1) can be replaced by a more general stochastic process X, and the linear Itô integral equation (1) by a different integral equation with respect to X. In this paper, the stochastic process X is called the *return process*, and the unique solution of an integral equation with respect to X is called the *stock price process* P.

[1]This research was partially supported by NSERC Canada grant number 203232-98 at York University and by Lithuanian State Science and Studies Foundation Grant K014.
[2]This research was partially supported by NSERC Canada grant number 203232-98 at York University.

The mainstream econometric analysis of continuous time financial models is to test different hypotheses about an integral equation describing a stock price process P, or to test various parameters of a distribution of a return process X (see e.g. Section 9.3 in Campbell, Lo and MacKinlay, 1997). In this paper we attempt to test the degree of roughness of a return process X. The legitimacy of such an endeavor is based on the fact that the support of a distribution of a stochastic processes is a particular class of functions. That is, a suitable class of functions contains almost all sample functions of a stochastic process. More specifically, if X is a regular enough stochastic process defined on a given probability space $(\Omega, \mathcal{F}, \Pr)$, then for a suitable class of functions F, a *sample function* $X(\omega) = \{X(t,\omega)\colon 0 \leq t \leq T\}$ belongs to F for almost all $\omega \in \Omega$. Often F can be taken as a proper subspace of the space of all continuous functions on $[0,T]$, or a proper subspace of the (Skorohod) space of all regulated and right continuous functions on $[0,T]$. For example, the support of the distribution of a standard Brownian motion is a subset of a class of functions having the order of Hölder continuity strictly bigger than $1/2$. However, the Hölder continuity is not applicable to characterize sample functions of a Lévy stochastic process without a Gaussian component, because almost every sample function of such a process is discontinuous. A simple example of a Lévy process is a symmetric α-stable stochastic process, or a $S\alpha S$ stochastic process, X_α with the exponent $\alpha \in (0,2]$. The case $\alpha = 2$ gives the only Gaussian component, a standard Brownian motion, that is, $X_2 = B$. In general, if X_α is a $S\alpha S$ stochastic process with $\alpha \in (0,2]$, then almost all sample functions of X_α have bounded p-variation for each $p > \alpha$, and have infinite p-variation if $0 < p \leq \alpha$. This fact was known since the time P. Lévy first introduced this process around the beginning of thirties.

To recall the property of boundedness of p-variation, let f be a real-valued function on an interval $[0,T]$. For a number $0 < p < \infty$, let

$$s_p(f;\kappa) := \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|^p,$$

where $\kappa = \{t_i\colon i = 0,\dots,n\}$ is a partition of $[0,T]$, that is $0 = t_0 < t_1 < \cdots < t_n = T$. The *p-variation* of f is defined by

$$v_p(f) := v_p(f;[0,T]) := \sup\{s_p(f;\kappa)\colon \kappa \text{ is a partition of } [0,T]\}.$$

The function f has bounded p-variation on $[0,T]$ if $v_p(f) < \infty$. By Hölder's inequality it follows that if $v_p(f) < \infty$ and $q > p$, then $v_q(f) < \infty$. The number

$$\upsilon(f;[0,T]) := \inf\{p > 0\colon v_p(f;[0,T]) < \infty\} = \sup\{p > 0\colon v_p(f;[0,T]) = \infty\}$$

is called the *p-variation index* of f. For a regular enough stochastic process X, $\upsilon(X)(\omega) := \upsilon(X(\cdot,\omega))$, $\omega \in \Omega$, is a random variable which we call the p-variation index of X. In fact, for all stochastic processes discussed in this paper, their p-variation indices are known to be constants and we seek to estimate these constants.

Now for a $S\alpha S$ stochastic process X_α, we can restate its sample regularity by saying that its p-variation index $\upsilon(X_\alpha) = \alpha$ almost surely. This is a special case of the following more general fact.

Example 1. Let X be a homogeneous Lévy stochastic process with the Lévy measure ν, which is a σ-finite Borel measure on $\mathbb{R} \setminus \{0\}$ such that

$$\int_{\mathbb{R}\setminus\{0\}} \min\{1, |x|^2\}\, \nu(dx) < +\infty.$$

The *Blumenthal-Getoor index* β_X of X is defined by

$$\beta_X := \inf\{\alpha > 0\colon \int_{\mathbb{R}\setminus\{0\}} \min\{1, |x|^\alpha\}\,\nu(dx) < +\infty\}.$$

Note that $0 < \beta_X \le 2$. If X has no Gaussian part, then for any $0 < T < \infty$,

$$\upsilon(X; [0,T]) = \beta_X, \qquad \text{almost surely.}$$

This follows from Theorems 4.1 and 4.2 of Blumenthal and Getoor (1961), and from Theorem 2 of Monroe (1972).

A stock price model having as a return a Lévy process without a Gaussian part is a common alternative to the Black-Scholes-Merton model. Another popular alternative is a fractional Brownian motion B_H with the Hurst exponent $H \in (0,1)$, where B_H with $H = 1/2$ is a standard Brownian motion. The p-variation index of a fractional Brownian motion $\upsilon(B_H; [0,T]) = 1/H$ almost surely. This is the special case of the following fact.

Example 2. Let $X = \{X(t)\colon t \ge 0\}$ be a Gaussian stochastic process with stationary increments and continuous in quadratic mean. Let σ_X be the incremental variance of X defined by $\sigma_X(u)^2 := E[(X(t+u) - X(t))^2]$ for $t, u \ge 0$. Let

$$\gamma_* := \inf\{\gamma > 0\colon \lim_{u\downarrow 0} \frac{u^\gamma}{\sigma_X(u)} = 0\} \quad \text{and} \quad \gamma^* := \sup\{\gamma > 0\colon \lim_{u\downarrow 0} \frac{u^\gamma}{\sigma_X(u)} = +\infty\}.$$

Then $0 \le \gamma_* \le \gamma^* \le +\infty$. If $\gamma_* = \gamma^*$, then we say that X has an *Orey index* $\gamma_X := \gamma_* = \gamma^*$. Furthermore, if X has an Orey index $\gamma_X \in (0,1)$, then for any $0 < T < +\infty$,

$$\upsilon(X; [0,T]) = 1/\gamma_X, \qquad \text{almost surely.} \tag{2}$$

This follows from the fact that almost all sample functions of X obey a uniform Hölder condition with exponent $\gamma < \gamma_X$ (see Section 9.4 of Cramer and Leadbetter, 1967) and from the inequality of Berman (1969) connecting the p-variation with the Fourier transform of local times of X. Relation (2) also follows from the characterization of the p-variation index for arbitrarily Gaussian processes due to Jain and Monrad (1983).

From the point of view of a statistical time series analysis, estimation of the p-variation index in the above two examples offer a new perspective to analyzing financial data sets. For instance, a symmetric α-stable process and a fractional Brownian motion with the Hurst exponent H, both have the same p-variation index in the case $\alpha = 1/H \in (1,2)$. However the latter has exponentially small tails, while the former has not even the second moment. These two examples are extensions of the Black-Scholes-Merton model (that is when $\alpha = 1/H = 2$) into two different directions. The Orey index, and so the p-variation index by relation (2), have already been estimated in the paper Norvaiša and Salopek (2000). They used two estimators based on the result of Gladyshev (1961). The estimators of the present paper can be applied under much less restrictive hypotheses about stock price returns, and helps to reconcile the two divergent directions of theoretical analyses of financial markets.

2 The oscillation η-summing index and related estimators

In this section we describe a method of estimating the p-variation index of a function based on existence of the metric entropy index (or the box-counting dimension) of its graph. Let f be a real-valued function defined on an interval $[0,T]$. Let $\eta = \{N_m\colon m \geq 1\}$ be a sequence of strictly increasing positive integers. With η one can associate a sequence $\{\lambda(m)\colon m \geq 1\}$ of partitions $\lambda(m) = \{iT/N_m\colon i = 0,\ldots,N_m\}$ of $[0,T]$ into subintervals $\Delta_{i,m} := [(i-1)T/N_m, iT/N_m]$, $i = 1,\ldots,N_m$, all having the same length T/N_m. For each $m \geq 1$, let

$$Q(f;\lambda(m)) = \sum_{i=1}^{N_m} \mathrm{Osc}(f;\Delta_{i,m}), \tag{3}$$

where for a subset $A \subset [0,T]$,

$$\mathrm{Osc}(f;A) := \sup\{|f(t)-f(s)|\colon s,t \in A\} = \sup_{t\in A} f(t) - \inf_{s\in A} f(s).$$

The sequence $Q_\eta(f) := \{Q(f;\lambda(m))\colon m \geq 1\}$ will be called the *oscillation η-summing* sequence. For a bounded non-constant function f on $[0,T]$, and any sequence η as above, let

$$\delta_\eta^-(f) := \liminf_{m\to\infty} \frac{\log Q(f;\lambda(m))/N_m}{\log(1/N_m)} \quad \text{and} \quad \delta_\eta^+(f) := \limsup_{m\to\infty} \frac{\log Q(f;\lambda(m))/N_m}{\log(1/N_m)}.$$

Then we have

$$0 \leq \delta_\eta^-(f) \leq \delta_\eta^+(f) \leq 1. \tag{4}$$

The lower bound follows from the bound $Q(f;\lambda(m))/N_m \leq \mathrm{Osc}(f;[0,T]) < \infty$, which is valid for each $m \geq 1$. The upper bound holds because $Q(f;\lambda(m)) \geq C/2 > 0$ for all sufficiently large $m \geq 1$. Indeed, if f is continuous, then $C = v_1(f;[0,T])$. Otherwise f has a jump at some $t \in [0,T]$, so that C can be taken to be a saltus at t if $t \notin \lambda(m)$ for all sufficiently large m, or C can be taken to be a one-sided non-zero saltus at t if $t \in \lambda(m)$ for infinitely many m. Instead of relation (4), a sharp one-sided bound is given by Lemma 13 in Appendix A.

If f has bounded variation, then for each η,

$$\delta_\eta^-(f) = \delta_\eta^+(f) = 1. \tag{5}$$

Indeed, since $Q(f;\lambda(m)) \leq v_1(f;[0,T]) < \infty$, we have $\log Q(f;\lambda(m))/N_m \geq \log v_1(f;[0,T])/N_m$ for all sufficiently large m. Thus $\delta_\eta^-(f) \geq 1$, and equalities in (5) follow from relation (4).

Definition 3. Let f be a non-constant real-valued function on $[0,T]$, and let $\eta = \{N_m\colon m \geq 1\}$ be a sequence of strictly increasing positive integers. If $\delta_\eta^-(f) = \delta_\eta^+(f)$, then we say that f has the *oscillation η-summing index* $\delta_\eta(f)$ and is defined by

$$\delta_\eta(f) := \delta_\eta^-(f) = \delta_\eta^+(f) = \lim_{m\to\infty} \frac{\log Q(f;\lambda(m))/N_m}{\log(1/N_m)}. \tag{6}$$

Next we give a sufficient condition for exitence of the oscillation η-summing index for any η. Let E be a nonempty bounded subset in a plane $\mathbb{R}^2$, and let $N(E;\epsilon)$, $\epsilon > 0$, be the minimum number of closed balls of diameter ϵ required to cover E. The *lower* and *upper metric entropy indices* of the set E are defined respectively by

$$\Delta^-(E) := \liminf_{\epsilon\downarrow 0} \frac{\log N(E;\epsilon)}{\log(1/\epsilon)} \quad \text{and} \quad \Delta^+(E) := \limsup_{\epsilon\downarrow 0} \frac{\log N(E;\epsilon)}{\log(1/\epsilon)}.$$

If $\Delta^-(E) = \Delta^+(E)$, then the common value denoted by $\Delta(E)$ is called the *metric entropy index* of the set E. In the actual calculations of the metric entropy index, it is often simpler to replace closed balls by squares (boxes) of a grid (cf. Lemma 9 below). Therefore in fractal analysis, $\Delta(E)$ is also known as the box-counting dimension.

The proof of the following theorem is given in Appendix A.

Theorem 4. *Let f be a regulated non-constant function on $[0,T]$ with the p-variation index $\upsilon(f)$. If the metric entropy index of the graph $gr(f)$ of f is defined and*

$$\Delta(gr(f)) = 2 - 1/(1 \vee \upsilon(f)), \tag{7}$$

then for any sequence η, f has the oscillation η-summing index

$$\delta_\eta(f) = 1/(1 \vee \upsilon(f)). \tag{8}$$

Essentially, relation (7) is the lower bound condition on the metric entropy index because the following always holds.

Proposition 5. *For a regulated function f on $[0,T]$, $\Delta^+(gr(f)) \le 2 - 1/(1 \vee \upsilon(f))$.*

The proof is similar to the proof of Theorem 4 and is also given in Appendix A.

The oscillation η-summing index is a slightly modified concept of a real box-counting method introduced by Carter, Cawley and Mauldin (1988). Independently, Dubuc, Quiniou, Roques-Carmes, Tricot and Zucker (1989) arrived at essentially the same notion, but called it the variation method. Both papers applied the new index to estimate the fractal dimension of several continuous functions whose dimension is known, and found the new algorithm superior over several other fractal dimension estimation methods (see also Section 6.2 in Cutler, 1993, for further discussion on this).

Oscillation η-summing estimators. Let f be a real-valued function defined on $[0,1]$, and let $\eta = \{N_m \colon m \ge 1\}$ be a sequence of strictly increasing positive integers. Let $\{u_1, \dots, u_N\} \subset [0,1]$ be a set of points such that for some integer M,

$$\cup_{m=1}^{M} \lambda(m) = \{u_1, \dots, u_N\}, \qquad \text{where } \lambda(m) := \{i/N_m \colon i = 0, \dots, N_m\}. \tag{9}$$

Given a finite set of values $\{f(u_1), \dots, f(u_N)\}$, we want to estimate the p-variation index $\upsilon(f)$. To achieve this, for each $m \in \{1, \dots, M\}$, let

$$Q(m) := \sum_{i=1}^{N_m} \Big[\max_{u_k \in \Delta_{i,m}} \{f(u_k)\} - \min_{u_k \in \Delta_{i,m}} \{f(u_k)\} \Big], \tag{10}$$

where $\Delta_{i,m} = [(i-1)/N_m, i/N_m]$. For large enough M, the finite set $\{Q(m) \colon m = 1, \dots, M\}$ may be considered as an approximation to the oscillation η-summing sequence $Q_\eta(f)$ defined by (3). For $m = 1, \dots, M$, let

$$r(m) := \frac{\log_2 1/N_m}{\log_2 Q(m)/N_m} = \frac{\log_2 N_m}{\log_2 N_m/Q(m)}. \tag{11}$$

Relation (8) suggests that the set $\{r(m) \colon m = 1, \dots, M\}$ may be used to estimate the p-variation index $\upsilon(f)$.

Definition 6. Let $\eta = \{N_m \colon m \geq 1\}$ be a sequence of strictly increasing positive integers, let $\{u_1, \ldots, u_N\} \subset [0,1]$ be such that (9) holds for some integer M, and let $\{f(u_1), \ldots, f(u_N)\}$ be a set of known values of a real-valued function f on $[0,1]$. We will say that $\widetilde{\upsilon}_\eta(f) := r(M)$ is the *naive oscillation η-summing estimator* of $1 \vee \upsilon(f)$ based on $\eta_M := \{N_m \colon 1 \leq m \leq M\}$. Letting $x_m := \log_2(N_m/Q(m))$ and $\bar{x} := M^{-1} \sum_{m=1}^M x_m$, we will say that

$$\widehat{\upsilon}_\eta(f) := \frac{\sum_{m=1}^M (x_m - \bar{x}) \log_2 N_m}{\sum_{m=1}^M (x_m - \bar{x})^2}$$

is the *OLS oscillation η-summing estimator* of $1 \vee \upsilon(f)$ based on $\eta_M := \{N_m \colon 1 \leq m \leq M\}$. The estimators $\widetilde{\upsilon}_\eta$ and $\widehat{\upsilon}_\eta$ will be called the OS estimators, and the estimation either by $\widetilde{\upsilon}_\eta$ or by $\widehat{\upsilon}_\eta$ will be called the OS estimation.

Relation (8) alone, if it holds for a function f and a sequence η, does not imply that the two estimators will converge to $\upsilon(f)$ as $N \to \infty$, and so as $M \to \infty$ by relation (9). If $\upsilon(f) < \infty$, then f is a regulated function on $[0,1]$, that is, there exist the limits $f(t+) := \lim_{u \downarrow t} f(u)$ for each $t \in [0,1)$ and $f(s-) := \lim_{u \uparrow s} f(u)$ for each $s \in (0,1]$. Assuming that f is regulated and either right- or left-continuous, then $\mathrm{Osc}(f; A)$ is the same as $\mathrm{Osc}(f; A \cap U)$ for a countable and dense subset $U \subset [0,1]$ and any subset $A \subset [0,1]$. For such a function f, one can show that the naive estimator $\widetilde{\upsilon}_\eta(f)$ will approach $\upsilon(f)$ as the set $\{u_1, \ldots, u_N\}$ will increase to $\cup_m \lambda(m)$. For sample functions of a stationary Gaussian stochastic process X, Hall and Wood (1993) showed that the two estimators corresponding to the reciprocal of relation (11) converge to $1/(1 \vee \upsilon(X))$, and they also calculated asymptotic bias and variance.

Oscillation η-summing index of stochastic processes. Here we show that the conditions of Theorem 4 hold for almost all sample functions of several important classes of stochastic processes. To this aim we use known results on Hausdorff-Besicovitch dimension of graphs of sample functions. Let $E \subset \mathbb{R}^2$ be a bounded set, and let $\mathrm{diam}(A)$ denote the diameter of a set $A \subset \mathbb{R}^2$. An ϵ-covering of E is a countable collection $\{E_k \colon k \geq 1\}$ of sets such that $E \subset \cup_k E_k$ and $\sup_k \mathrm{diam}(E_k) \leq \epsilon$. For $s > 0$, the Hausdorff s-measure of E is defined by

$$\mathcal{H}^s(E) := \lim_{\epsilon \downarrow 0} \inf \Big\{ \sum_{k \geq 1} (\mathrm{diam}(E_k))^s \colon \{E_k \colon k \geq 1\} \text{ is an } \epsilon\text{-covering of } E \Big\}.$$

Given $E \subset \mathbb{R}^2$, the function $s \mapsto \mathcal{H}^s(E)$ is nonincreasing. In fact, there is a critical value s_c such that $\mathcal{H}^s(E) = \infty$ for $s < s_c$ and $\mathcal{H}^s(E) = 0$ for $s > s_c$. This critical value s_c is called the *Hausdorff-Besicovitch dimension* and is denoted by $\dim_{HB}(E)$. That is,

$$\dim_{HB}(E) := \inf \{s > 0 \colon \mathcal{H}^s(E) = 0\} = \sup\{s > 0 \colon \mathcal{H}^s(E) = \infty\}.$$

A relation between the lower metric entropy index $\Delta^-(E)$ of E and the Hausdorff-Besicovitch dimension of E is given by the following result.

Lemma 7. *For a bounded subset $E \subset \mathbb{R}^2$, $\dim_{HB}(E) \leq \Delta^-(E)$.*

Proof. Let $E \subset \mathbb{R}^2$ be bounded, and let $s > \Delta^-(E)$. Since

$$\Delta^-(E) = \sup \{\alpha > 0 \colon \lim_{\epsilon \downarrow 0} N(E; \epsilon) \epsilon^\alpha = +\infty\},$$

we have that $\lim_{\epsilon\downarrow 0} N(E;\epsilon)\epsilon^s = 0$. Thus, for each $\epsilon > 0$, there exists an ϵ-covering of E by $N(E;\epsilon)$ balls of equal diameter ϵ. Hence

$$\inf\Big\{\sum_{k\geq 1}(\mathrm{diam}(E_k))^s\colon \{E_k\colon k\geq 1\} \text{ is an } \epsilon\text{-covering of } E\Big\} \leq N(E;\epsilon)\epsilon^s,$$

which gives $\mathcal{H}^s(E) \leq \liminf_{\epsilon\downarrow 0} N(E;\epsilon)\epsilon^s = 0$. Thus $\dim_{HB}(E) \leq s$, and so $\dim_{HB}(E) \leq \Delta^-(E)$, proving the lemma. Q.E.D.

Let $\mathrm{gr}(X)$ be the graph of a regulated sample function of a stochastic process X. By Proposition 5 and by the preceding lemma, we have

$$\dim_{HB}(\mathrm{gr}(X)) \leq \Delta^-(\mathrm{gr}(X)) \leq \Delta^+(\mathrm{gr}(X)) \leq 2 - 1/(1\vee \upsilon(X)). \tag{12}$$

In fact, the left side is equal to the right side almost surely for several classes of stochastic processes. For example, let X_α be a symmetric α-stable process for some $0<\alpha\leq 2$. Then by Theorem B of Blumenthal and Getoor (1962), for almost all sample functions of X_α, we have

$$\dim_{HB}(\mathrm{gr}(X_\alpha)) = 2 - \frac{1}{1\vee\alpha} = 2 - \frac{1}{1\vee \upsilon(X_\alpha)}.$$

For another example, let X be a stochastic process of Example 2 having the Orey index $\gamma_X \in (0,1)$. Then by Theorem 1 of Orey (1970), for almost all sample functions of X, we have

$$\dim_{HB}(\mathrm{gr}(X)) = 2 - \gamma_X = 2 - \frac{1}{\upsilon(X)}.$$

For these stochastic processes, relation (12) yields that assumption (7) of Theorem 4 is satisfied for almost every sample function, and so we have the following result.

Corollary 8. *Let $X = \{X(t)\colon 0\leq t\leq 1\}$ be a stochastic process, and let $\eta = \{N_m\colon m\geq 1\}$ be a sequence of strictly increasing positive integers. The relation*

$$\delta_\eta(X) \equiv \lim_{m\to\infty}\frac{\log Q(X;\lambda(m))/N_m}{\log(1/N_m)} = \frac{1}{1\vee \upsilon(X)}$$

holds for almost all sample functions of X provided either (a) or (b) holds, where

(*a*) *X is a symmetric α-stable process for some $\alpha\in(0,2]$;*

(*b*) *X is a mean zero Gaussian stochastic process with stationary increments continuous in quadratic mean and such that the Orey index γ_X exists.*

Similar results for more general processes other than (*a*) and (*b*) of Corollary 8 can be respectively found in Pruitt and Taylor (1969, Section 8) and in Kôno (1986).

3 Simulated symmetric α-stable process

In this section we carry out a simulation study of small-sample properties of the OS estimators from Definition 6. To this aim, we simulate a symmetric α-stable process for several values of the exponent α, which is equal to its p-variation index. Using repeated samples we calculate the bias, the standard deviation and the mean square error for the two estimators.

Table 1: Properties of 100 samples of estimates $\widetilde{\upsilon}_\eta(X_\alpha)$ based on $\{2^m\colon 1 \le m \le 14\}$

α	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
$\overline{\alpha}$	1.278	1.346	1.406	1.473	1.556	1.632	1.715	1.798	1.883	1.967	2.051
$\overline{\alpha}-\alpha$	.2783	.2464	.2064	.1727	.1563	.1322	.1149	.0984	.0832	.0667	.0510
SD	.0867	.0864	.0542	.0285	.0264	.0147	.0089	.0063	.0059	.0040	.0025
MSE	.0849	.0681	.0455	.0306	.0251	.0177	.0133	.0097	.0070	.0045	.0026

Table 2: Properties of 100 samples of estimates $\widehat{\upsilon}_\eta(X_\alpha)$ based on $\{2^m\colon 1 \le m \le 14\}$

α	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
$\overline{\alpha}$	1.158	1.197	1.269	1.337	1.383	1.464	1.529	1.598	1.654	1.711	1.767
$\overline{\alpha}-\alpha$	.1582	.0974	.0693	.0375	-.0172	-.0362	-.0712	-.1022	-.1459	-.1887	-.2329
SD	.0726	.0785	.0902	.0804	.0964	.0794	.0791	.0786	.0691	.0555	.0435
MSE	.0302	.0156	.0129	.0078	.0095	.0076	.0113	.0166	.0260	.0387	.0560

Simulating $S\alpha S$ process. Let $X_\alpha = \{X_\alpha(t)\colon t \ge 0\}$ be a symmetric α-stable stochastic process with the exponent $\alpha \in (0, 2]$. As stated in the introduction, the p-variation index of X_α is given by $\upsilon(X_\alpha) = \alpha$ almost surely. Since the OS estimators do not capture the values of the p-variation index below 1, we restrict our study to estimating the exponent $\alpha \in [1, 2]$. To simulate a sample function of a $S\alpha S$ process X_α, we generate a set $\{\xi_i\colon i = 1, \dots, n\}$ of symmetric α-stable pseudo-random variables and use the central limit theorem to get an approximation $\widetilde{X}_\alpha$ of X_α, where

$$\widetilde{X}_\alpha(t) = \frac{1}{n^{1/\alpha}} \sum_{i=1}^{[nt]} \xi_i, \qquad 0 \le t \le 1,$$

and $[r]$ denotes the integer part of r. By the central limit theorem, the distribution of $\widetilde{X}_\alpha$ on the Skorohod space $D[0, 1]$ converges weakly to the distribution of X_α as $n \to \infty$. We take $n = 2^{14}$. To generate a symmetric α-stable random variable ξ, we use the results of Chambers et al. (1976) (see also Section 4.6 in Zolotarev 1986). That is, in the sense of equality in distribution, we have

$$\xi = \frac{\sin \alpha U}{(\cos U)^{1/\alpha}} \cdot \Big(\frac{\cos(U - \alpha U)}{E}\Big)^{(1-\alpha)/\alpha} \quad \text{if } \alpha \ne 1, \qquad \text{and} \qquad \xi = \tan U \quad \text{if } \alpha = 1,$$

where the random variable U has the uniform distribution on $[-\pi/2, \pi/2]$, and the random variable E, which is independent of U, has the standard exponential distribution. All calculations are done using the computing system *Mathematica*.

OS estimators. Let $\eta = \{2^m\colon m \ge 1\}$. We simulate a sample function $\widetilde{X}_\alpha$ at $2^{14} + 1$ equally spaced points $\{u_1, \dots, u_N\} = \{i2^{-14}\colon i = 0, \dots, 2^{14}\}$. For each $1 \le m \le 14$, let

$$Q(m) := \sum_{i=1}^{2^m} \Big[\max_{u_k \in \Delta_{i,m}} \{\widetilde{X}_\alpha(u_k)\} - \min_{u_k \in \Delta_{i,m}} \{\widetilde{X}_\alpha(u_k)\} \Big],$$

where $\Delta_{i,m} := [(i-1)2^{-m}, i2^{-m}]$. Then define $r(m)$, $m = 1, \dots, 14$, by relation (11) with $N_m = 2^m$. We use the OS estimators based on $\eta_{14} = \{2^m\colon 1 \le m \le 14\}$. Thus by Definition 6,

the naive oscillation η-summing estimator $\widetilde{\upsilon}_\eta(X_\alpha) = r(14)$, and the OLS oscillation η-summing estimator

$$\widehat{\upsilon}_\eta(X_\alpha) = \frac{\sum_{m=1}^{14}(x_m - \bar{x})m}{\sum_{m=1}^{14}(x_m - \bar{x})^2},$$

where $x_m = \log_2(2^m/Q(m))$ and $\bar{x} = 14^{-1}\sum_{m=1}^{14} x_m$.

Monte-Carlo study. For 11 different values of $\alpha \in \{1.0, 1.1, \ldots, 2.0\}$, we simulate a vector $\{\widetilde{X}_\alpha(k/2^M) : k = 0, \ldots, 2^M\}$ of values of a sample function of X_α with $M = 14$. By the preceding paragraph, we have two estimators $\widehat{\alpha}$ of α: the naive oscillation η-summing estimator $\widetilde{\upsilon}_\eta(X_\alpha)$, and the OLS oscillation η-summing estimator $\widehat{\upsilon}_\eta(X_\alpha)$, both based on $\{2^m\colon 1 \le m \le 14\}$. We repeat this procedure $K = 100$ times to obtain the estimates $\widehat{\alpha}_1, \ldots \widehat{\alpha}_K$ of α for each of the two cases. Then we calculate:

- The estimated expected value $\overline{\alpha} := (\sum_{i=1}^K \widehat{\alpha}_i)/K$;
- The bias $\overline{\alpha} - \alpha$;
- The estimated standard deviation $SD := \sqrt{\sum_{i=1}^K (\widehat{\alpha}_i - \overline{\alpha})^2/(K-1)}$;
- The estimated mean square error $MSE := \sum_{i=1}^K (\widehat{\alpha}_i - \alpha)^2/K$.

The estimation results are presented in Tables 1 and 2. Next is a qualitative description of the performance of the two OS estimators.

Bias The estimation results show a different behavior of the bias for the two estimators. The naive estimator $\widetilde{\upsilon}_\eta(X_\alpha)$ display monotonically decreasing positive bias when α values increase from 1 to 2. While the bias of the OLS estimator $\widehat{\upsilon}_\eta(X_\alpha)$ monotonically decrease from a positive bias for $\alpha < 1.4$ to a negative bias for $\alpha \ge 1.4$, and the minimal absolute bias is achieved when $\alpha = 1.4$.

SD The estimated standard deviation is also different for the two estimators. SD values for the naive estimator $\widetilde{\upsilon}_\eta(X_\alpha)$ monotonically decrease when α values increase from 1 to 2, and are quite small when α is close to 2. While SD values for the OLS estimator $\widehat{\upsilon}_\eta(X_\alpha)$ remain similar and somewhat larger than for the naive estimator; only a little improvement one can notice when α is close to 2.

MSE The estimated mean square error remain different for the two estimators. MSE values for the naive estimator $\widetilde{\upsilon}_\eta(X_\alpha)$ decrease steady when α values increase from 1 to 2, while MSE values for the OLS estimator $\widetilde{\upsilon}_\eta(X_\alpha)$ is smallest when $\alpha = 1.5$, and are increasing for all other values of α.

In conclusion the results show a distinction between the two OS estimators: the OLS estimator $\widetilde{\upsilon}_\eta(X_\alpha)$ display better performance for values $1.4 \le \alpha \le 1.6$, while the naive estimator $\widetilde{\upsilon}_\eta(X_\alpha)$ behaves best for α values close to 2. This Monte-Carlo study was extended to sample functions based on a larger number of points, i.e. $\{\widetilde{X}_\alpha(k/2^M) : k = 0, \ldots, 2^M\}$ with $M \in \{15, 16, 17\}$. The results from the increased sample size show the same qualitative behavior as before, but with increased accuracy (see Norvaiša and Salopek, 2000b).

To the best of our knowledge, the two OS estimators provide the first attempt to estimate the exponent α of a $S\alpha S$ process from a sample function. Recently Crovella and Taqqu (1999) introduced a new method to estimate the exponent α of a $S\alpha S$ random variable.

Table 3: Properties of 100 samples of estimates $\widetilde{\upsilon}_\eta(B_H)$ based on $\{2^m\colon 1 \le m \le 14\}$

$h = 1/H$	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$\overline{h}$	1.132	1.304	1.474	1.640	1.802	1.961	2.117	2.270	2.420	2.566
$\overline{h} - h$	-.0683	-.0956	-.1262	-.1603	-.1979	-.2390	-.2826	-.3296	-.3804	-.4339
SD	.0039	.0018	.0013	.0020	.0021	.0027	.0026	.0035	.0042	.0042
MSE	.0047	.0092	.0159	.0257	.0392	.0571	.0799	.1086	.1447	.1883

Table 4: Properties of 100 samples of estimates $\widehat{\upsilon}_\eta(B_H)$ based on $\{2^m\colon 1 \le m \le 14\}$

$h = 1/H$	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$\overline{h}$	1.186	1.340	1.481	1.614	1.712	1.821	1.891	1.971	2.043	2.095
$\overline{h} - h$	-.0143	-.0598	-.1185	-.1864	-.2879	-.3795	-.5089	-.6294	-.7572	-.9053
SD	.0490	.0466	.0499	.0469	.0465	.0474	.0399	.0447	.0494	.0471
MSE	.0026	.0057	.0165	.0369	.0850	.1462	.2606	.3982	.5757	.8217

4 Simulated fractional Brownian motion

Here we perform a Monte Carlo study like in the preceding section, with a fractional Brownian motion. A fractional Brownian motion $\{B_H(t)\colon t \ge 0\}$ with the Hurst exponent $H \in (0,1)$ is a Gaussian stochastic process with stationary increments having the covariance function

$$E[B_H(t)B_H(s)] = \frac{1}{2}\Big[t^{2H} + s^{2H} - |t-s|^{2H}\Big], \qquad \text{for } t, s \ge 0$$

and $B_H(0) = 0$ almost surely. Since its incremental variance $\sigma_{B_H}(u) = u^H$, the Orey index γ_{B_H} exists and is equal to the Hurst exponent H (cf. Example 2). Thus by relation (2), B_H has the p-variation index $\upsilon(B_H) = 1/H$ almost surely. To simulate a fractional Brownian motion we use the program of Maeder (1995) written in *Mathematica*.

OS and G estimation. In this section we apply four estimators of the p-variation index. As before, the two OS estimators from Definition 6 will be used to estimate $h := 1/H$. Moreover, we invoke the two estimators of the Orey index introduced in Norvaiša and Salopek (2000), and which will be called *G estimation*, which is short for the Gladyshev estimation. More specifically, let $\eta = \{2^m\colon m \ge 1\}$. We simulate a sample function $\widetilde{B}_H$ at $2^{14}+1$ equally spaced points $\{u_1, \dots, u_N\} = \{i2^{-14}\colon i = 0, \dots, 2^{14}\}$. For each $1 \le m \le 14$, let

$$Q(m) := \sum_{i=1}^{2^m} \Big[\max_{u_k \in \Delta_{i,m}} \{\widetilde{B}_H(u_k)\} - \min_{u_k \in \Delta_{i,m}} \{\widetilde{B}_H(u_k)\}\Big],$$

where $\Delta_{i,m} := [(i-1)2^{-m}, i2^{-m}]$. Then define $r(m)$, $m = 1, \dots, 14$, by (11) with $N_m = 2^m$. Therefore one can use estimators based on $\eta_{14} = \{2^m\colon 1 \le m \le 14\}$. Thus by Definition 6, the naive oscillation η-summing estimator $\widetilde{\upsilon}_\eta(B_H) = r(14)$, and the OLS oscillation η-summing estimator

$$\widehat{\upsilon}_\eta(B_H) = \frac{\sum_{m=1}^{14}(x_m - \bar{x})m}{\sum_{m=1}^{14}(x_m - \bar{x})^2},$$

Table 5: Properties of 100 samples of estimates $\widetilde{1/\gamma_\eta}(B_H)$ based on $\{2^m\colon 1 \le m \le 14\}$

$h = 1/H$	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$\overline{h}$	1.164	1.348	1.530	1.709	1.887	2.062	2.236	2.407	2.576	2.742
$\overline{h} - h$	-.0364	-.0524	-.0704	-.0907	-.1133	-.1383	-.1645	-.1928	-.2245	-.2580
SD	.0039	.0020	.0013	.0023	.0022	.0031	.0030	.0039	.0044	.0048
MSE	.0013	.0027	.0050	.0082	.0128	.0191	.0271	.0372	.0504	.0666

Table 6: Properties of 100 samples of estimates $\widehat{1/\gamma_\eta}(B_H)$ based on $\{2^m\colon 1 \le m \le 14\}$

$h = 1/H$	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$\overline{h}$	1.236	1.429	1.628	1.818	1.987	2.184	2.325	2.483	2.645	2.790
$\overline{h} - h$	.0363	.0292	.0281	.0185	-.0131	-.0161	-.0751	-.1171	-.1548	-.2100
SD	.0705	.0870	.1209	.1241	.1418	.1689	.1425	.2459	.2650	.2343
MSE	.0062	.0084	.0153	.0156	.0201	.0285	.0258	.0736	.0935	.0985

where $x_m = \log_2(2^m/Q(m))$ and $\bar{x} = 14^{-1}\sum_{m=1}^{14} x_m$.

To recall the G estimation, again let $\eta = \{2^m\colon m \ge 1\}$, and let $\widetilde{B}_H$ be a sample function given by its values at $2^{14}+1$ equally spaced points $\{i2^{-14}\colon i = 0, \ldots, 2^{14}\}$. For each $1 \le m \le 14$, let

$$s_2(m) := \sum_{i=1}^{2^m} \Big[\widetilde{B}_H(i/N_m) - \widetilde{B}_H((i-1)/N_m)\Big]^2 \quad \text{and} \quad r(m) := \frac{\log 2^{-m}}{\log \sqrt{s_2(m)2^{-m}}}.$$

The *naive Gladyshev estimator* of the p-variation index is defined by $\widetilde{1/\gamma_\eta}(B_H) = r(14)$, and the OLS Gladyshev estimator is defined by

$$\widehat{1/\gamma_\eta}(B_H) = \frac{\sum_{m=1}^{14}(x_m - \bar{x})m}{\sum_{m=1}^{14}(x_m - \bar{x})^2},$$

where $x_m = \log_2 \sqrt{2^m/s_2(m)}$ and $\bar{x} = 14^{-1}\sum_{m=1}^{14} x_m$.

Monte-Carlo study. For 10 different values of $H \in \{0.83 \approx 1.2^{-1}, 0.71 \approx 1.4^{-1}, \ldots, 0.33 \approx 3.0^{-1}\}$, we simulate the vector $\{B_H(i2^{-m}) : i = 0, \ldots, 2^m\}$ with $m = 14$, and calculate the four estimators. This gives us four different estimates $\widehat{h}$ of $h = 1/H$. We repeat this procedure $K = 100$ times to obtain the estimates $\widehat{h}_1, \ldots \widehat{h}_K$ of h for each of the four cases. Then we calculate:

- The estimated expected value $\overline{h} := (\sum_{i=1}^K \widehat{h}_i)/K$;
- The bias $\overline{h} - h$;
- The estimated standard deviation $SD := \sqrt{\sum_{i=1}^K (\widehat{h}_i - \overline{h})^2/(K-1)}$;
- The estimated mean square error $MSE := \sum_{i=1}^K (\widehat{h}_i - h)^2/K$.

Table 7: Properties of 100 samples of estimates $\widetilde{1/\upsilon}_\eta(B_H)$ based on $\{2^m\colon 1 \le m \le 14\}$

H	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\overline{H}$	.1578	.2572	.3566	.4559	.5549	.6539	.7526	.8511	.9496
$\overline{H}-H$	.0578	.0572	.0566	.0559	.0549	.0539	.0526	.0511	.0496
SD	.0006	.0007	.0006	.0007	.0006	.0007	.0010	.0020	.0101
MSE	.0033	.0033	.0032	.0231	.0030	.0029	.0027	.0026	.0026

Table 8: Properties of 100 samples of estimates $\widehat{1/\upsilon}_\eta(B_H)$ based on $\{2^m\colon 1 \le m \le 14\}$

H	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\overline{H}$	.3471	.3932	.4481	.5109	.5802	.6508	.7320	.8146	.8962
$\overline{H}-H$	.2471	.1932	.1481	.1109	.0802	.0508	.0320	.0146	-.0038
SD	.0069	.0081	.0113	.0141	.0175	.0235	.0276	.0364	.0449
MSE	.0611	.0374	.0221	.0125	.0067	.0031	.0018	.0015	.0020

The estimation results are presented in Tables 3 - 6. First one can compare the OS estimation results of $\alpha \in [1,2]$ (Tables 1 and 2) for a $S\alpha S$ process, and the OS estimation results of $h = 1/H \in (1,2]$ (columns $h = 1.2, 1.4, 1.6, 1.8, 2.0$ of Tables 3 and 4) for a fractional Brownian motion with the Hurst exponent H. The accuracy of the naive OS estimator for the two processes is similar. There is only some differences in the character of monotonicity along different values of parameters. The same holds for the OLS OS estimator for the two processes.

Now if we look at columns $h \in \{2.0, 2.2, \ldots, 3.0\}$ of Tables 3 and 4, it is clear that the OS estimates of h are very poor in this case. This is so since estimation errors appear in the denominator of the relation (11). For example, if H and $\widehat{H}$ both are small, then the left side of the relation

$$\frac{1}{H} - \frac{1}{\widehat{H}} = (\widehat{H} - H)\frac{1}{H\widehat{H}} \tag{13}$$

can be relatively large as compared to $\widehat{H} - H$. The same remark applies to the G estimates in Tables 5 and 6. To show that this is so we applied the OS estimator to evaluate the Hurst exponent directly using the relations (2) and (8). This means that in our earlier estimation formulas we need just to interchange the numerator and the denominator in (11). The estimation results are presented in Tables 7 and 8, where $\widehat{H}$, the bias, SD and MSE are defined as before with h replaced by H. In Tables 3 - 8, the estimation results for $H = 1/2$, and so for $h = 2$, are all based on the same set of 100 simulated sample functions, which can be used to verify the effect of the above relation 13. Also, the results of the OS estimation of H in Tables 7 and 8 can be compared with the results of the G estimation of H in Tables 1 and 3 of Norvaiša and Salopek (2000). The naive estimators corresponding to the OS and G estimations show very similar properties. As far as the OLS estimators concern, for small vaules of H, the bias and MSE of the OS estimation are larger than the bias and MSE of G estimation. However, for the same values of H, the standard deviation of the OS estimation is smaller than the standard deviation of the G estimation. In sum the two estimation methods OS and G show similar results when applied to estimate the Hurst exponent of a fractional Brownian motion.

5 Financial data analysis

In this section, we analyze the financial data set provided by Olsen & Associates. It is the high-frequency data set HFDF96, which consists of 25 different foreign exchange spot rates, 4 spot metal rates, and 2 series of stock indices. This data set was recorded from 1 Jan 1996 GMT to 31 Dec 1996 GMT. Each set has 17568 entries recorded at half hour intervals. The same financial data set was studied in Norvaiša and Salopek (2000) using the two estimators of the Orey index based on the result of Gladyshev (1961).

Returns. First notice that returns in continuous time and discrete time financial models are treated slightly differently. A return in a discrete time model is a function $\widehat{R}$ defined on a lattice $t \in \{0, 1, \ldots, T\}$ with values being a suitable transform of a pair $\{P(t-1), P(t)\}$, where P is a stock price process. A return in a continuous time model is a function R defined on $[0, T]$ so that $R(t) - R(t-1) = \widehat{R}(t)$ and $R(0) = 0$ for all $t \in \{1, \ldots, T\}$. This gives a $1-1$ correspondence between continuous time returns used in this paper and the usual discrete time returns (see Section 2.1 in Norvaiša, 2000a, for further details).

Given a historical data set $\{d_0, \ldots, d_K\}$ of values of a financial asset, let P be a function defined on $[0, 1]$ with values d_k at $u_k := k/K$ for $k = 0, \ldots, K$. Usually in econometric literature, returns are log transforms of the price. Thus in continuous time models, this corresponds to assuming that $P(t) = P(0)\exp\{X(t)\}$, $0 \le t \le 1$, for some return process X which is to be analyzed. Alternatively, the price process P can be a solution of an integral equation, which is not a simple exponential. If a stochastic process X has the quadratic variation along the sequence of partitions $\lambda = \{\lambda(m)\colon m \ge 1\}$ defined by (9), and the price process P is a solution of a linear integral equation with respect to X (such as the Black-Scholes-Merton model (1), with B replaced by X), then P also has the quadratic variation along the sequence λ and the process X can be recovered by the relation

$$X(t) = R_{net}(P)(t) := \lim_{m\to\infty} \sum_{i=1}^{N_m} [P(t \wedge i/N_m) - P(t \wedge (i-1)/N_m)]/P(t \wedge (i-1)/N_m),$$

for $0 \le t \le 1$. Here the quadratic variation is understood in the sense of Fölmer (1981), which is further developed in Norvaiša (2000b). The above process X will be called the *net return* of P, which is analogous to discrete time simple net returns. The *log return* X of the price process P is defined by $X(t) = R_{log}(P)(t) := \log[P(t)/P(0)]$, $0 \le t \le 1$. The difference between the two returns is

$$R_{net}(P)(t) - R_{log}(P)(t) = \frac{1}{2}\int_0^t \frac{d[P]^c}{P^2} - \sum_{(0,t]} \Big[\log\frac{P}{P_-} - \frac{\Delta^- P}{P_-}\Big] - \sum_{[0,t)} \Big[\log\frac{P_+}{P} - \frac{\Delta^+ P}{P}\Big] \tag{14}$$

for $0 \le t \le 1$, where $[P]^c$ is a continuous part of the quadratic variation of P. Since P has the quadratic variation along the sequence λ, the difference $R_{net}(P) - R_{log}(P)$ given by (14) has bounded variation. Thus the p-variation indices of the two returns are equal provided both are not less than 1. A discussion of the difference $R_{net}(P) - R_{log}(P)$ when P is the geometric Brownian motion or P is a model for USD/JPY exchange rates can be found on pages 362 and 366 of Norvaiša and Salopek (2000).

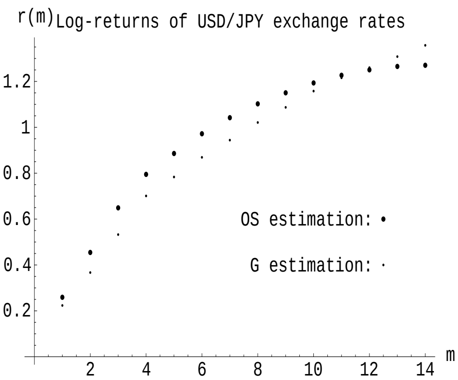

Figure 1: Naive estimators: $\widetilde{\upsilon}_\eta = 1.27$ and $\widetilde{1/\gamma}_\eta = 1.357$ based on $\{2^m\colon m = 1, \ldots, 14\}$.

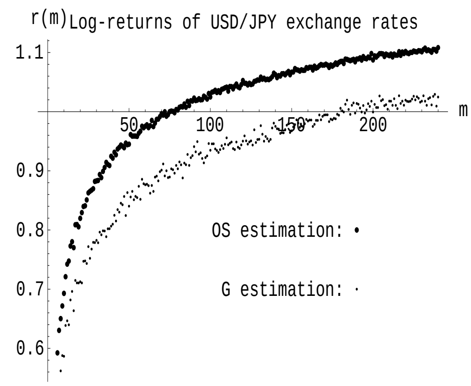

Figure 2: Naive estimators: $\widetilde{\upsilon}_\eta = 1.108$ and $\widetilde{1/\gamma}_\eta = 1.025$ based on $\{m\colon m = 1, \ldots, 240\}$.

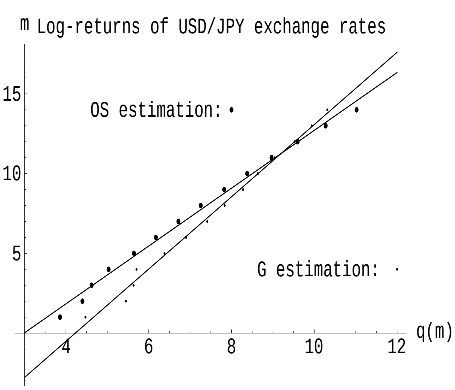

Figure 3: OLS estimators: $\widehat{\upsilon}_\eta = 1.814$ and $\widehat{1/\gamma}_\eta = 2.266$ based on $\{2^m\colon m = 1, \ldots, 14\}$.

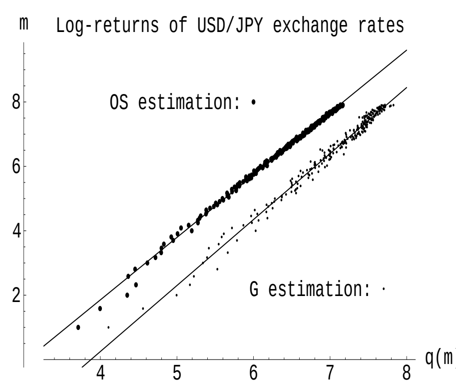

Figure 4: OLS estimators: $\widehat{\upsilon}_\eta = 1.939$ and $\widehat{1/\gamma}_\eta = 2.049$ based on $\{m\colon m = 1, \ldots, 240\}$.

OS and G estimation. To estimate the p-variation index of a return, the oscillation η-summing estimators from Definition 6 will be used when the sequence $\eta = \{N_m\colon m \geq 1\}$ is given by $N_m = 2^m$ and $N_m = m$ for integers $m \geq 1$. A comparison will be made with the results of Norvaiša and Salopek (2000) by using their estimators based on the result of Gladyshev (1961). As before, the oscillation η-summing estimation is called the *OS estimation*, and the estimation as in Norvaiša and Salopek (2000) is called the *G estimation*.

Given a price process P at points $u_k = k/K$, $k = 0, \ldots, K$, find the maximal integer M such that

$$\cup_{m=1}^{M} \lambda(m) \subset \{u_0, \ldots, u_K\}, \qquad \text{where } \lambda(m) = \{i/N_m\colon i = 0, \ldots, N_m\}. \tag{15}$$

Notice that in the case $N_m = m$, the sequence $\{\lambda(m)\colon m \geq 1\}$ is *not* nested. Since for HFDF96 data set $K = 17568$, $M = 14$ when $N_m = 2^m$ and $M = 240$ when $N_m = m$.

Let $\eta = \{N_m\colon m \geq 1\}$ be a sequence of strictly increasing positive integers, and let M be such that relation (15) holds. In the case $X = R_{net}(P)$ or $X = R_{log}(P)$, for each $1 \leq m \leq M$, let

$$Q(m) := Q(X; \lambda(m)) = \sum_{i=1}^{N_m} \Big[\max_{u_k \in \Delta_{i,m}} \{X(u_k)\} - \min_{u_k \in \Delta_{i,m}} \{X(u_k)\} \Big],$$

where $\Delta_{i,m} := [(i-1)/N_m, i/N_m]$. Also for each $1 \leq m \leq M$, let

$$q(m) := \log_2(N_m/Q(m)) \quad \text{and} \quad r(m) := \frac{\log(1/N_m)}{\log Q(m)/N_m} = \frac{\log_2 N_m}{q(m)}. \tag{16}$$

By Definition 6, the naive oscillation η-summing estimator $\widetilde{\upsilon}_\eta := \widetilde{\upsilon}_\eta(X) = r(M)$. The OLS oscillation η-summing estimator $\widehat{\upsilon}_\eta$ is defined by

$$\widehat{\upsilon}_\eta := \widehat{\upsilon}_\eta(X) = \frac{\sum_{m=1}^{M}(x_m - \bar{x}) \log_2 N_m}{\sum_{m=1}^{M}(x_m - \bar{x})^2}, \tag{17}$$

where $x_m = q(m) = \log_2(N_m/Q(m))$ and $\bar{x} = M^{-1} \sum_{m=1}^{M} x_m$.

Turning to the G estimation, again let $\eta = \{N_m\colon m \geq 1\}$ be a sequence of strictly increasing positive integers, and let M be such that relation (15) holds. In the case $X = R_{net}(P)$ or $X = R_{log}(P)$, for each $1 \leq m \leq M$, let

$$s_2(m) := \sum_{i=1}^{N_m} \Big[X(i/N_m) - X((i-1)/N_m) \Big]^2.$$

In relations (16) and (17), for each $1 \leq m \leq M$, replacing $N_m/Q(m)$ by $\sqrt{N_m/s_2(m)}$, let

$$q(m) := \log_2 \sqrt{N_m/s_2(m)} \quad \text{and} \quad r(m) := \frac{\log(1/N_m)}{\log \sqrt{s_2(m)/N_m}} = \frac{\log_2 N_m}{q(m)}.$$

Then the *naive Gladyshev estimator* of the p-variation index is $\widetilde{1/\gamma}_\eta := \widetilde{1/\gamma}_\eta(X) = r(M)$. The OLS Gladyshev estimator $\widehat{1/\gamma}_\eta$ is defined by

$$\widehat{1/\gamma}_\eta := \widehat{1/\gamma}_\eta(X) = \frac{\sum_{m=1}^{M}(x_m - \bar{x}) \log_2 N_m}{\sum_{m=1}^{M}(x_m - \bar{x})^2},$$

where $x_m = q(m) = \log_2 \sqrt{N_m/s_2(m)}$ and $\bar{x} = M^{-1} \sum_{m=1}^{M} x_m$.

Table 9: OS and G estimation of returns of exchanges rates based on $\{2^m\colon 1 \leq m \leq 14\}$.

Currency	net-returns				log-returns			
	$\widetilde{\upsilon}_\eta$	$\widehat{\upsilon}_\eta$	$\widetilde{1/\gamma}_\eta$	$\widehat{1/\gamma}_\eta$	$\widetilde{\upsilon}_\eta$	$\widehat{\upsilon}_\eta$	$\widetilde{1/\gamma}_\eta$	$\widehat{1/\gamma}_\eta$
AUD/USD	1.2452	1.7617	1.3563	2.2436	1.2452	1.7664	1.3563	2.2663
CAD/USD	1.1493	1.8362	1.2362	2.7382	1.1493	1.8357	1.2362	2.7422
DEM/ESP	1.1504	1.7197	1.4043	2.5684	1.1504	1.7274	1.4042	3.4243
DEM/FIM	1.2124	1.8537	1.3566	2.4970	1.2124	1.8571	1.3566	2.5047
DEM/ITL	1.2420	1.7619	1.3673	2.0673	1.2420	1.7552	1.3673	2.0496
DEM/JPY	1.2671	1.8045	1.3661	2.4714	1.2671	1.8034	1.3662	2.4770
DEM/SEK	1.2483	1.7601	1.3767	2.2612	1.2483	1.7558	1.3767	2.2359
GBP/DEM	1.2325	1.6459	1.3380	1.9566	1.2325	1.6493	1.3380	1.9678
GBP/USD	1.2398	1.7708	1.3334	2.1373	1.2398	1.7725	1.3334	2.1466
USD/BEF	1.2615	1.7083	1.4745	2.4220	1.2615	1.7176	1.4745	2.5022
USD/CHF	1.2796	1.7000	1.3902	2.0817	1.2796	1.7044	1.3902	2.1001
USD/DEM	1.2432	1.7646	1.3416	2.2111	1.2432	1.7673	1.3416	2.2294
USD/DKK	1.2896	1.5574	1.7142	2.2990	1.2894	1.6096	1.7102	2.6473
USD/ESP	1.3574	1.9885	1.5331	2.5367	1.3574	2.0155	1.5331	2.7526
USD/FIM	1.3208	1.8583	1.4520	2.2860	1.3208	1.8682	1.4520	2.3226
USD/FRF	1.2467	1.7718	1.3643	2.2542	1.2467	1.7766	1.3643	2.2811
USD/ITL	1.2977	2.1020	1.4115	3.0934	1.2977	2.0863	1.4115	2.9658
USD/NLG	1.2591	1.7570	1.3815	2.2647	1.2591	1.7622	1.3816	2.2937
USD/SEK	1.3263	1.9542	1.4453	2.5304	1.3263	1.9606	1.4453	2.5575
USD/XEU	1.2450	1.8503	1.3594	2.5097	1.2450	1.8539	1.3594	2.5423
USD/JPY	1.2703	1.8104	1.3573	2.2434	1.2703	1.8143	1.3573	2.2662
USD/MYR	1.1121	1.6715	1.3957	2.9137	1.1121	1.6905	1.3956	3.0156
USD/SGD	1.1377	1.8821	1.2584	3.2629	1.1377	1.8794	1.2584	3.2128
USD/ZAR	1.2377	1.5460	1.4400	1.9238	1.2377	1.5516	1.4401	1.9449

Estimation results. Estimation results for the data set HFDF96 are given by Tables 9, 10 and 11. More specifically, the tables contain the estimated p-variation indices for the two returns of the bid price associated with the nearest prior datum. We picked the log-returns of USD/JPY exchange rates to illustrate by Figures 1 - 4 a difference between the OS and G estimation results in more detail. Figures 2 and 4 show the estimation results based on a sequence $\eta = \{N_m = m\colon m \geq 1\}$ truncated at $M = 240$. The associated sequence of partitions $\{\lambda(m)\colon m \geq 1\}$ in this case is not nested, and technical calculations in this case are more complex. Figures 1 and 3 show the estimation results based on dyadic partitions, which are used for the rest of results. Because the conditions of Theorem 4 are more general as compared to the conditions of the main result of Gladyshev (1961), the results of the OS estimation are more reliable than the results of G estimation. This is also seen from the Figures 2 and 4. The columns $\widetilde{\upsilon}_\eta$ and $\widehat{\upsilon}_\eta$ of Tables 9, 10 and 11 suggest that estimated p-variation indices of the returns of the financial data are more likely to belong to the interval $(1, 2)$. These columns are essentially the same for net-returns as well as for log-returns, which would be in agreement if a stock price has the quadratic variation along a nested sequence of dyadic partitions and its p-variation index is

Table 10: OS and G estimation of returns of metal rates based on $\{2^m\colon 1 \le m \le 14\}$.

Metal	net-returns				log-returns			
	$\widetilde{\upsilon}_\eta$	$\widehat{\upsilon}_\eta$	$\widetilde{1/\gamma}_\eta$	$\widehat{1/\gamma}_\eta$	$\widetilde{\upsilon}_\eta$	$\widehat{\upsilon}_\eta$	$\widetilde{1/\gamma}_\eta$	$\widehat{1/\gamma}_\eta$
Gold	1.2190	1.7249	1.3488	2.2808	1.2190	1.7216	1.3489	2.2700
Silver	1.4097	1.8482	1.6227	2.6066	1.4097	1.8263	1.6229	2.5728
Palladium	1.3371	1.6650	1.5481	2.3535	1.3371	1.6603	1.5480	2.3390
Platinum	1.2358	1.6472	1.4134	2.2933	1.2358	1.6433	1.4135	2.2790

Table 11: OS and G estimation of returns of stock indices based on $\{2^m\colon 1 \le m \le 14\}$.

Index	net-returns				log-returns			
	$\widetilde{\upsilon}_\eta$	$\widehat{\upsilon}_\eta$	$\widetilde{1/\gamma}_\eta$	$\widehat{1/\gamma}_\eta$	$\widetilde{\upsilon}_\eta$	$\widehat{\upsilon}_\eta$	$\widetilde{1/\gamma}_\eta$	$\widehat{1/\gamma}_\eta$
SP 500	1.1685	1.3723	1.3726	1.9085	1.1685	1.3734	1.3727	1.9162
DOW JONES	1.1806	1.3909	1.3937	1.9035	1.1806	1.3928	1.3938	1.9133

not less than 1. Due to the pattern exhibit by the results of Section 4 when estimating $1/H$, it is unlikely that a fractional Brownian motion B_H would give a satisfactory fit to the financial data for some values of H. However, as seen in Figures 1 - 4, a more general process from Example 2 might be used to model the HFDF96 data sets. Especially this concerns a modelling of stock indices (see Figure 11). Recall that the G estimation is reliable when applied to sample functions of stochastic processes having a suitable relationship between a sample function behavior and an asymptotic behavior of the incremental variance. The estimation results do not reject the hypothesis that returns may be modeled by a Lévy process. To be more specific about a degree of data fitting, we need a theoretical asymptotic analysis of both, the OS and G estimations, which is not available at this writing.

Related results and techniques. The OS and G estimators are based on properties of a function similar to a kind of self-similarity property with respect to shrinking time intervals, sometimes refered to as a fractal or empirical scaling law. It is natural that a high-frequency data have already been used to detect such laws if exist. The work of Müller et. al. (1995) addresses this question and provide some preliminary findings in their analysis of a high-frequency FX data collected from raw data vendors such as Reuters, Knight-Rider and Telerate. Another related work of Mandelbrot (1997) have already been discussed in Section 4.3 of Norvaiša and Salopek (2000).

6 Discussion

Parameter estimation of a financial model is a typical econometric analysis task. What is atypical in the preceding analysis is a generality of the underlying financial model. This model applies far beyond of limits imposed by the semimartingale theory, and its outline can be found in Bick and Willinger (1994), and Norvaiša (2000a). The p-variation index considered as a parameter of a model is defined for any function. Its estimate provides a grade for each concrete continuous time model of a price process P governed by an exponential, or by a linear integral equation

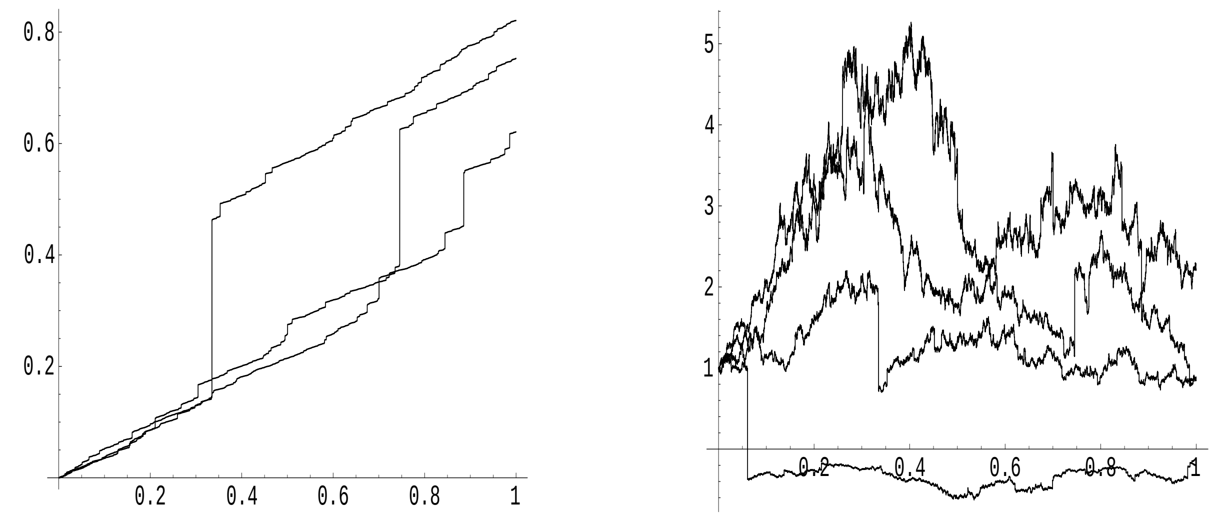

Figure 5: Four trajectories of $R_{net}(P) - R_{log}(P)$ (left) defined by (14) corresponding to the Dolean exponential $P = \mathcal{E}(X_\alpha)$ (right) given by (19), where X_α is a SαS process with $\alpha = 1.7$. Vertical lines join jumps.

having the indefinite integral

$$P(t) = \lim_{\kappa} \prod_{i=1}^{n} (1 + X(t_i \wedge t) - X(t_{i-1} \wedge t)), \qquad 0 < t \leq 1, \tag{18}$$

as its unique solution, where the limit is understood either in the sense of refinements of partitions $\kappa = \{t_i\colon i = 0, \ldots, n\}$ of $[0,1]$, or in the more general sense along a fixed sequence of nested partitions. The linear Itô stochastic integral equation with respect to a semimartingale X is one such example of a continuous time model. A further generality of the underlying financial model could be achieved once the net-returns are modified so as to reverse a solution of a non-linear integral equation.

We stress the importance of the notion of a return because it provides a two direction link between theory and practice. A financial model without the notion of a return is just an exercise in theory building. As we noted earlier, the results of the preceding section show that the estimated p-variation indices for the calculated net-returns and log-returns are almost the same. Here we discuss the difference between the net- and log-returns $R_{net}(P) - R_{log}(P)$ given by (14) for a simulated price process P. Suppose that P is the Dolean exponential $\mathcal{E}(X_\alpha)$ of a symmetric α-stable process X_α:

$$\mathcal{E}(X_\alpha)(t) := \exp\{X_\alpha(t) - X_\alpha(0)\} \prod_{(0,t]} (1 + \Delta^- X_\alpha) e^{-\Delta^- X_\alpha}, \qquad 0 < t \leq 1. \tag{19}$$

That is, P is defined by (18) with X replaced by X_α. Then the continuous part of the quadratic variation $[P]^c \equiv 0$, and so the simulation gives the remaining sums in (14) as shown in Figure 5.

To simulate the Dolean exponential $\mathcal{E}(X_\alpha)$ we use its representation as the indefinite product integral (18), proved by Dudley and Norvaiša (1999, Corollary 5.23 in Part II). Theoretically, the Dolean exponential composed with the net-return gives $X_\alpha = R_{net}(P)$. Thus the maximal discrepancy $d := \sup_t |\widetilde{X}_\alpha(t) - \widetilde{R}_{net}(\widetilde{P})(t)|$ between simulated versions of the two sides should be

small if simulation is accurate enough. Indeed, we get that the discrepency $d = 3.1 \times 10^{-14}$ holds uniformly for all trajectories in the left Figure 5. Notice that one of the four trajectories breaks because its Dolean exponential jumps to the negative side, and so its logarithm is undefined.

Appendix A

This section contains proofs of Theorem 4 and Proposition 5. The proofs consist of combining a few relations between the metric entropy index, the oscillation η-summing index and the p-variation index, and will be discussed first.

As noted earlier, in the actual calculations of the metric entropy it is often simpler to replace closed balls by squares (boxes), which leads to the box-counting dimension (see Section 6.1 in Cutler, 1993). First we modify the box-counting method for sets which arise when connecting a graph of a discontinuous function. Throughout this section, f is a regulated function defined on a closed interval $[0, T]$. This means that for each $t \in (0, T]$, there exists a limit $f(t-) := \lim_{u \uparrow t} f(u)$, and for each $t \in [0, T)$, there exists a limit $f(t+) := \lim_{u \downarrow t} f(u)$. For $t \in (0, T]$, let $V_-(f(t))$ be the interval connecting the points $(t, f(t-))$ and $(t, f(t))$, and for $t \in [0, T)$, let $V_+(f(t))$ be the interval connecting the points $(t, f(t))$ and $(t, f(t+))$. On a plane $\mathbb{R}^2$ with the Cartesian coordinates (t, x), for each $0 \le u < v \le T$, let

$$G_f([u,v]) := (u, V_+(f(u))) \bigcup \Big(\bigcup_{u<t<v} (t, V_+(f(t))) \bigcup (t, V_-(f(t))) \Big) \bigcup (v, V_-(f(v))) \subset \mathbb{R}^2, \quad (20)$$

where $(t, A(t)) := \{(t, x) \colon x \in A(t)\}$. Let $G_f := G_f([0, T])$, and let $\mathrm{gr}(f)$ be the graph of the function f. It is clear that $\mathrm{gr}(f) \subset G_f$, and the equality holds between the two sets if and only if f is continuous on $[0, T]$. For $\epsilon > 0$, a *grid* $\mathcal{C}_\epsilon$ of side length ϵ covering a bounded set E, is called a collection of disjoint squares of equal side length ϵ whose union is a square containing E. We can also assume that each grid $\mathcal{C}_\epsilon$ intersects the origin of the Cartesian coordinates, and its squares have sides parallel to the coordinates axes. The usual box-counting method counts all squares of the grid $\mathcal{C}_\epsilon$ which have nonempty intersection with E. We need a special counting rule when a vertical segment of the set $E = G_f$ is a part of a vertical line of a grid $\mathcal{C}_\epsilon$. Let $\{t_i \colon i = 0, \ldots, N_\epsilon\}$ be the partition of $[0, T]$ induced by intersecting a grid $\mathcal{C}_\epsilon$ with the t-axes, and for each $i = 1, \ldots, N_\epsilon$, let $n_i(\epsilon)$ be the number of squares of the grid $\mathcal{C}_\epsilon$ contained in the strip $[t_{i-1}, t_i] \times \mathbb{R}$ which have a nonempty intersection with $G_f([t_{i-1}, t_i])$. For each $\epsilon > 0$, we then define

$$M(G_f; \epsilon) := \sum_{i=1}^{N_\epsilon} n_i(\epsilon).$$

This definition avoids a double covering of certain parts of vertical segments, and it reduces to the usual definition if f is continuous. The present extension allows us to apply Theorem 4 to sample functions of Lévy processes, which may have left-side discontinuities.

Next statement shows the usefulness of the box-counting method.

Lemma 9. *Let f be a regulated function on $[0, T]$. For lower and upper metric entropy indices, we have*

$$\Delta^-(G_f) = \liminf_{\epsilon \downarrow 0} \frac{\log M(G_f; \epsilon)}{\log(1/\epsilon)} \quad \text{and} \quad \Delta^+(G_f) = \limsup_{\epsilon \downarrow 0} \frac{\log M(G_f; \epsilon)}{\log(1/\epsilon)}.$$

Proof. For $\epsilon > 0$, let $\mathcal{C}_\epsilon$ be a grid of side length ϵ covering G_f. Every square of side ϵ is included in the ball of diameter $\epsilon\sqrt{2}$, which, in turn, is included in at most 9 squares of the grid $\mathcal{C}_\epsilon$. Thus

$$N(G_f; \epsilon\sqrt{2}) \leq M(G_f; \epsilon) \leq 9N(G_f; \epsilon\sqrt{2}),$$

proving the lemma. Q.E.D.

Next we look at replacing a limit as $\epsilon \downarrow 0$ by a limit along a countable sequence. Let $\{\epsilon_m : m \geq 1\}$ be a sequence of positive real numbers strictly decreasing to 0, and for each $m \geq 1$, let $\Delta_m := (\epsilon_{m+1}, \epsilon_m]$. For a family $\{A_\epsilon : \epsilon > 0\}$ of numbers, we have

$$\limsup_{\epsilon \downarrow 0} A_\epsilon = \limsup_{m\to\infty} \sup_{\epsilon \in \Delta_m} A_\epsilon \geq \limsup_{m\to\infty} A_{\epsilon_m}$$

and

$$\liminf_{\epsilon \downarrow 0} A_\epsilon = \liminf_{m\to\infty} \inf_{\epsilon \in \Delta_m} A_\epsilon \leq \liminf_{m\to\infty} A_{\epsilon_m}.$$

Thus by the preceding lemma, we have that

$$\Delta^-(G_f) \leq \liminf_{m\to\infty} \frac{\log M(G_f; \epsilon_m)}{\log(1/\epsilon_m)} \quad \text{and} \quad \limsup_{m\to\infty} \frac{\log M(G_f; \epsilon_m)}{\log(1/\epsilon_m)} \leq \Delta^+(G_f). \tag{21}$$

Lemma 10. *Let f be a regulated function on $[0, T]$, and let $\{\epsilon_m : m \geq 1\}$ be a sequence of positive real numbers strictly decreasing to 0 so that*

$$\lim_{m\to\infty} \frac{\log \epsilon_m}{\log \epsilon_{m+1}} = 1. \tag{22}$$

Then

$$\left\{ \begin{array}{l} \liminf_{\epsilon \downarrow 0} \\ \limsup_{\epsilon \uparrow 0} \end{array} \right\} \frac{\log M(G_f; \epsilon)}{\log(1/\epsilon)} = \left\{ \begin{array}{l} \liminf_{m\to\infty} \\ \limsup_{m\to\infty} \end{array} \right\} \frac{\log M(G_f; \epsilon_m)}{\log(1/\epsilon_m)}. \tag{23}$$

Proof. It is enough to prove the reverse inequalities in relations (21). For $0 < \epsilon \leq \epsilon_1$, let $m \geq 1$ be such that $\epsilon \in \Delta_m = (\epsilon_{m+1}, \epsilon_m]$. Since a square with a side length ϵ is contained in at most four squares with a side length ϵ_m, we have $M(G_f; \epsilon_m) \leq 4M(G_f; \epsilon)$. Similarly, $M(G_f; \epsilon) \leq 4M(G_f; \epsilon_{m+1})$. Thus for each $\epsilon \in \Delta_m$,

$$\frac{\log \epsilon_m}{\log \epsilon_{m+1}} \Big(\frac{\log M(G_f; \epsilon_m) - \log 4}{\log 1/\epsilon_m} \Big) \leq \frac{\log M(G_f; \epsilon)}{\log(1/\epsilon)} \leq \frac{\log \epsilon_{m+1}}{\log \epsilon_m} \Big(\frac{\log M(G_f; \epsilon_{m+1}) + \log 4}{\log 1/\epsilon_{m+1}} \Big).$$

The conclusion now follows by the assumption (22). Q.E.D.

Now recall relation (3) defining the oscillation η-summing sequence $\{Q(f; \lambda(m)) : m \geq 1\}$.

Lemma 11. *Let f be a regulated non-constant function on $[0, T]$. For any sequence $\{N_m : m \geq 1\}$ of strictly increasing positive integers, we have*

$$\left\{ \begin{array}{l} \liminf_{m\to\infty} \\ \limsup_{m\to\infty} \end{array} \right\} \frac{\log M(G_f; T/N_m)}{\log N_m} = \left\{ \begin{array}{l} \liminf_{m\to\infty} \\ \limsup_{m\to\infty} \end{array} \right\} \frac{\log N_m Q(f; \lambda(m))}{\log N_m}. \tag{24}$$

Proof. Let $\{N_m\colon m \geq 1\}$ be a sequence of strictly increasing positive integers. For each $m \geq 1$, let A_m and B_m be the terms under the limit signs in equality (24) in the given order. It is enough to prove that

$$\liminf_m A_m \leq \liminf_m B_m, \qquad \limsup_m A_m \leq \limsup_m B_m \tag{25}$$

and

$$\liminf_m B_m \leq \liminf_m A_m, \qquad \limsup_m B_m \leq \limsup_m A_m. \tag{26}$$

Suppose that the set G_f is on a plane with the Cartesian coordinates (t, x). For each integer $m \geq 1$, let $\mathcal{C}_m$ be the grid of side length T/N_m intersecting the origin and with sides parallel to the coordinate axes. Thus $\mathcal{C}_m$ intersects with the t-axes at each point of the partition $\lambda(m) = \{t_i^m = iT/N_m\colon i = 0, \ldots, N_m\}$. Let $m \geq 1$. For $i = 1, \ldots, N_m$, let $n_i = n_i(m)$ be the number of squares of the grid $\mathcal{C}_m$ contained in the strip $\Delta_{i,m} \times \mathbb{R}$ and covering the set $G_f(\Delta_{i,m})$ defined by relation (20), where $\Delta_{i,m} = [t_{i-1}^m, t_i^m]$. By the definition of $G_f(\Delta_{i,m})$,

$$(N_m/T)\mathrm{Osc}(f; \Delta_{i,m}) \leq n_i \leq 2 + (N_m/T)\mathrm{Osc}(f; \Delta_{i,m}),$$

for each $i = 1, \ldots, N_m$. Summing over all indices i, it follows that the inequalities

$$M(G_f; T/N_m) - 2N_m \leq (N_m/T)Q(f; \lambda(m)) \leq M(G_f; T/N_m) \tag{27}$$

hold for each $m \geq 1$. By the second inequality in display (27), we have that inequalities (26). To prove inequalities (25), first suppose that $\liminf_m B_m = 1$. Then by relation (4), $\limsup_m B_m = 1$, and inequalities (25) follow because $\liminf_m A_m \geq 1$. Now suppose that $\liminf_m B_m > 1$. In that case, inequalities (25) follow from the first inequality in display (27) and from the following two relations:

$$\liminf_{m\to\infty} \frac{\log a_m}{\log b_m} = \sup\{\alpha > 0\colon \lim_{m\to\infty} b_m^{-\alpha} a_m = +\infty\} = \sup\{\alpha > 0\colon \inf_{m\geq 1} b_m^{-\alpha} a_m > 0\}$$

and

$$\limsup_{m\to\infty} \frac{\log a_m}{\log b_m} = \inf\{\alpha > 0\colon \lim_{m\to\infty} b_m^{-\alpha} a_m = 0\} = \inf\{\alpha > 0\colon \sup_{m\geq 1} b_m^{-\alpha} a_m < +\infty\}, \tag{28}$$

valid for any two sequences $\{a_m\colon m \geq 1\}$ and $\{b_m\colon m \geq 1\}$ of positive numbers such that $\lim_{m\to\infty} b_m = +\infty$. Q.E.D.

Combining Lemmas 9, 10 and 11, it follows that the following statement holds.

Corollary 12. *Let f be a regulated non-constant function on $[0, T]$, and let $\{N_m\colon m \geq 1\}$ be a sequence of strictly increasing positive integers such that*

$$\lim_{m\to\infty} \frac{\log N_m}{\log N_{m+1}} = 1. \tag{29}$$

Then

$$\Delta(G_f) = \lim_{\epsilon\downarrow 0} \frac{\log M(G_f; \epsilon)}{\log(1/\epsilon)} = \lim_{m\to\infty} \frac{\log N_m Q(f; \lambda(m))}{\log N_m},$$

provided that at least one of the three limits exists and is finite.

Next is the final step in a chain of inequalities used to prove Theorem 4.

Lemma 13. *Let f be a regulated function on $[0,T]$, and let $\eta = \{N_m\colon m \ge 1\}$ be a sequence of strictly increasing positive integers. Then*

$$\Big(2-\delta_\eta^+(f)=\Big) \qquad \limsup_{m\to\infty} \frac{\log N_m Q(f;\lambda(m))}{\log N_m} \le 2-\frac{1}{1\vee v(f)}. \tag{30}$$

Proof. If $v(f)<1$ then f has bounded 1-variation, and so (30) holds by relation (5). If $v(f)=+\infty$, then (30) holds because $\delta_\eta^+(f)\ge 0$ by relation (4). Thus one can assume that $1\le v(f)<+\infty$. Let $v_p(f;[0,T])<\infty$ for some $1\le p<\infty$. We claim that for each integer $m\ge 1$,

$$N_m^{\frac{1}{p}-1} Q(f;\lambda(m)) \le v_p(f;[0,T])^{1/p}. \tag{31}$$

Indeed, if $p=1$ then $Q(f;\lambda(m))\le v_1(f;[0,T])$. Suppose that $p>1$ and $m\ge 1$. Since for each $i=1,\dots,N_m$, $\mathrm{Osc}(f;\Delta_{i,m})\le v_p(f;\Delta_{i,m})^{1/p}$, by Hölder's inequality, we have

$$\begin{aligned} Q(f;\lambda(m)) &\le \sum_{i=1}^{N_m} v_p(f;\Delta_{i,m})^{1/p} \le N_m^{1-1/p}\Big(\sum_{i=1}^{N_m} v_p(f;\Delta_{i,m})\Big)^{1/p} \\ \text{by subadditivity of } v_p \quad &\le N_m^{1-1/p} v_p(f;[0,T])^{1/p}, \end{aligned}$$

proving relation (31). If $p>v(f)\ge 1$, then by (31),

$$\sup_{m\ge 1} N_m^{-(2-1/p)}[N_m Q(f;\lambda(m))] \le v_p(f;[0,T])^{1/p} < +\infty.$$

Thus relation (30) follows by the relation (28), proving the lemma. Q.E.D.

Now we are ready to complete the proofs.

Proof of Theorem 4. Let $\eta=\{N_m\colon m\ge 1\}$ be a sequence of strictly increasing positive integers. We have to prove that $\delta_\eta(f)=1/(1\vee v(f))$. By Definition 3, this will be done once we will show that

$$\lim_{m\to\infty} \frac{\log N_m Q(f;\lambda(m))}{\log N_m} = 2-\frac{1}{1\vee v(f)}. \tag{32}$$

Since the graph $\mathrm{gr}(f)\subset G_f$, we have

$$\begin{aligned} \Delta^-(\mathrm{gr}(f)) &\le \Delta^-(G_f) = \liminf_{\epsilon\downarrow 0} \frac{\log M(G_f;\epsilon)}{\log(1/\epsilon)} \overset{\text{by (21)}}{\le} \liminf_{m\to\infty} \frac{\log M(G_f;1/N_m)}{\log N_m} \\ \text{by (24)} \quad &= \liminf_{m\to\infty} \frac{\log N_m Q(f;\lambda(m))}{\log N_m} \le \limsup_{m\to\infty} \frac{\log N_m Q(f;\lambda(m))}{\log N_m} \\ \text{by (30)} \quad &\le 2-\frac{1}{1\vee v(f)}. \end{aligned}$$

Since the left and right sides are equal by assumption (7), relation (32) holds, proving Theorem 4. Q.E.D.

Proof of Proposition 5. If f is constant then the conclusion clearly holds, and so we can assume that f is non-constant. Let $\eta=\{N_m\colon m\ge 1\}$ be a sequence of strictly increasing

positive integers such that relation (29) holds. Since the graph $\mathrm{gr}(f) \subset G_f$, we have

$$\begin{aligned} \Delta^+(\mathrm{gr}(f)) &= \limsup_{\epsilon \downarrow 0} \frac{\log M(G_f;\epsilon)}{\log(1/\epsilon)} \overset{\text{by (23)}}{=} \limsup_{m\to\infty} \frac{\log M(G_f;1/N_m)}{\log N_m} \\ \text{by (24)} &= \limsup_{m\to\infty} \frac{\log N_m Q(f;\lambda(m))}{\log N_m} \overset{\text{by (30)}}{\leq} 2 - \frac{1}{1 \vee \upsilon(f)}, \end{aligned}$$

proving Proposition 5. Q.E.D.

References

S. M. Berman, "Harmonic analysis of local times and sample functions of Gaussian processes," *Trans. Amer. Math. Soc.* vol. 143 pp. 269-281, 1969.

A. Bick and W. Willinger, "Dynamic spanning without probabilities," *Stoch. Proc. Appl.* vol. 50 pp. 349-374, 1994.

R. M. Blumenthal and R. K. Getoor, "Sample functions of stochastic processes with stationary independent increments," *J. Math. and Mech.* vol. 10 pp. 493-516, 1961.

R. M. Blumenthal and R. K. Getoor, "The dimension of the set of zeros and the graph of a symmetric stable process," *Illinois J. Math.* vol. 6 pp. 308-316, 1962.

J. Y. Campbell, A. W. Lo and A. C. MacKinlay, *The econometrics of financial markets*, Princeton University Press: New Jersey, 1997.

P. H. Carter, R. Cawley and R. D. Mauldin, "Mathematics of dimension measurement for graphs of functions," in *Fractal Aspects of Materials: Disordered Systems*, Eds. D. A. Weitz, L. Sander and B. Mandelbrot, Material Research Society, US, 1988, pp. 183-186.

J. M. Chambers, C. L. Mallows and B. W. Stuck, "A method for simulating stable random variables," *J. Amer. Statist. Assoc.* vol. 71 no. 354 pp. 340-344, 1976.

H. Cramer and M. R. Leadbetter, *Stationary and related stochastic processes*, Wiley: New York, 1967.

M. E. Crovella and M. S. Taqqu, "Estimating the heavy tail index from scaling properties," *Meth. and Comp. in Appl. Probab.* vol. 1 pp. 55-79, 1999.

C. D. Cutler, "A review of the theory and estimation of fractal dimension," in *Dimension Estimation and Models*, Ed. H. Tong., World Scientific, Singapore, 1993, pp. 1-107.

B. Dubuc, J. F. Quiniou, C. Roques-Carmes, C. Tricot and S. W. Zucker, "Evaluating the fractal dimension of profiles," *Physical Review A* vol. 39 pp. 1500-1512, 1989.

R. M. Dudley and R. Norvaiša, *Differentiability of Six Operators on Nonsmooth Functions and p-Variation*, Lect. Notes in Math., vol. 1703, Springer: Berlin, 1999

H. Fölmer, "Calcul d'Itô sans probabilités," in *Séminaire de Probabilités* XV, Eds. J. Azéma and M. Yor, Lect. Notes Math. vol. 850, Springer, 1981, pp. 143-150.

E. G. Gladyshev, "A new limit theorem for stochastic processes with Gaussian increments," *Theor. Probability Appl.* vol. 6 pp. 52-61, 1961.

P. Hall and A. Wood, "On the performance of box-counting estimators of fractal dimension," *Biometrika*, vol. 80 pp. 256-252, 1993.

N. C. Jain and D. Monrad, "Gaussian measures in B_p," *Ann. Probab.* vol. 11 pp. 46-57, 1983.

N. Kôno, "Hausdorff dimension of sample paths for self-similar processes," in *Dependence in Probability and Statistics*, Eds., E. Eberlein and M. S. Taqqu, Birkhäuser, Boston, 1986, pp. 109-117.

R. E. Maeder, “Fractional Brownian motion,” *Mathematica J* vol. 6 pp. 38-48, 1995.

B. B. Mandelbrot, *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Selecta volume E, Springer: New York, 1997.

I. Monroe, “On the γ-variation of processes with stationary independent increments,” *Ann. Math. Statist.* vol. 43 pp. 1213-1220, 1972.

U. A. Müller, M. M. Dacorogna, R. D. Davé, O. V. Pictet, R. B. Olsen and J. R. Ward, “Fractals and intrinsic time - a chalenge to econometricians,” Internal document UAM.1993-08-16, Olsen & Associates, Switzerland, 1995. http://www.olsen.ch

R. Norvaiša, “Modelling of stock price changes: A real analysis approach,” *Finance and Stochastics* vol. 4 pp. 343-369, 2000a.

R. Norvaiša, “Quadratic variation, p-variation and integration with applications to stock price modelling,” 2000b. (in preparation)

R. Norvaiša and D. M. Salopek, “Estimating the Orey index of a Gaussian stochastic process with stationary increments: An application to financial data set,” in *Stochastic Models*, Proc. Int. Conf., Ottawa, Canada June 10-13, 1998, Eds. L. G. Gorostiza and B. G. Ivanoff. Canadian Math. Soc., Conference Proceedings, vol. 26 2000, pp. 353-374.

R. Norvaiša and D. M. Salopek, “Supplement to the present paper, containing codes in Mathematica,” 2000b. (available from the authors upon request)

S. Orey, “Gaussian sample functions and the Hausdorff dimension of level crossings,” *Z. Wahrsch. verw. Gebiete* vol. 15 pp. 249-256, 1970.

W. E. Pruit and S. J. Taylor, “Sample path properties of processes with stable components,” *Z. Wahrsch. verw. Geb.* vol. 12 pp. 267-289, 1969.

V. M. Zolotarev, *One-dimensional Stable Distributions*, Translations of Mathematical Monographs, vol. 65, American Mathematical Society, 1986.